Large Language Models as Generalizable Policies for Embodied Tasks

Abstract

We show that large language models (LLMs) can be adapted to be generalizable policies for embodied visual tasks. Our approach, called Large LAnguage model Reinforcement Learning Policy (LLaRP), adapts a pre-trained frozen LLM to take as input text instructions and visual egocentric observations and output actions directly in the environment. Using reinforcement learning, we train LLaRP to see and act solely through environmental interactions. We show that LLaRP is robust to complex paraphrasings of task instructions and can generalize to new tasks that require novel optimal behavior. In particular, on 1,000 unseen tasks it achieves 42% success rate, 1.7x the success rate of other common learned baselines or zero-shot applications of LLMs. Finally, to aid the community in studying language conditioned, massively multi-task, embodied AI problems we release a novel benchmark, Language Rearrangement, consisting of 150,000 training and 1,000 testing tasks for language-conditioned rearrangement. Video examples of LLaRP in unseen Language Rearrangement instructions are at https://llm-rl.github.io.